A Practical Test of the Air Density Equation in Standards Laboratories at Differing Altitude

R. M. Schoonover, R. S. Davis, R. G. Driver and V. E. Bower

National Measurement Laboratory, National Bureau of Standards, Washington, D.C. 20234

A number of weighings of kilogram artifacts have been completed at sites of differing altitude. The artifacts and altitude difference were chosen to amplify the role of the necessary buoyancy corrections and thereby to uncover systematic errors in those corrections as they are usually applied. Small systematic effects were discovered but these are not explainable by buoyancy errors. Rather, we suggest their source is a lack of thermal equilibrium between the artifacts and the balance chamber.

Key words: Air density equation; buoyancy correction; double substitution weighing; uncertainty in weighing.

1. Introduction

In 1975 results were published of a series of measurements undertaken by the National Bureau of Standards (NBS) of the mass of aluminum and tantalum artifacts as determined by comparison against standards of stainless steel [1].[1] The paper reported inconsistencies which seemed to be correlated with barometric pressure. The stated magnitude of the unexpected effect is 1 mg in 1 kg over a pressure range from 0.5 to 2 atmospheres for objects having a volume difference of 200 cm³. The sign of the effect was not reported in [1]. The inconsistencies or "anomalies" as they were termed were observed between laboratories near sea level and those at an altitude of $\sim$ 1600 m. Quantitative results of these measurements are not given in the paper. An examination of the original data, however, shows that an aluminum kilogram (density $\sim$ 2.8 g cm⁻³) was measured to be 830 μg lighter compared to a stainless steel standard (density $\sim$ 7.8 g cm⁻³) at the higher altitude than at sea level. The tantalum kilogram (density $\sim$ 16.6 g cm⁻³), on the other hand, was found to be 275 μg heavier than at sea level.

A conclusion of [1] is that buoyant forces on objects placed in air are incorrectly accounted for by the usual means of computing ϱ, the density of air, from an equation whose input parameters include barometric pressure, temperature, relative humidity, and sometimes, CO_2 fraction.

Recently, Jones [2] has published a careful reformulation of the air density equation. He concludes that, using state-of-the-art measurements of pressure, temperature and rela-

tive humidity, the following relative uncertainties are to be expected in ϱ: 300 ppm (parts per million) random, 200 ppm systematic at a level corresponding to one standard deviation.

Koch, Davis and Bower [3] have intercompared two objects of different density to determine their mass difference both in vacuo and in air. From these measurements, they can test Jones' air density equation. The agreement is well within their experimental uncertainty of 600 ppm in ϱ.

In an effort to reconcile these measurements, which are consistent with Jones' air density equation, with those summarized in [1], the following experiment was undertaken. A series of weighings at NBS, Gaithersburg, was made with a selection of kilogram artifacts. Similar measurements with the same artifacts were also carried out at Sandia Laboratories, Albuquerque. The NBS, Gaithersburg, laboratories are near sea level while Sandia is $\sim$ 1600 m above sea level. The artifacts chosen included the aluminum and tantalum kilograms used in [1] as well as several other weights designed to elucidate surface effects. A great deal of care was taken to tie measurements of pressure, temperature and relative humidity directly to primary standards.

A very brief review of the principles involved in these measurements may be useful. Consider the comparison of two kilograms of nominally equal mass. Let M and V be the mass and volume of the standard, let M_x and V_x be the mass and volume of the unknown, and let ϱ be the density of air inside the balance case. The balance responds to forces. Under equilibrium conditions, there are two forces which must be considered: gravitational, and buoyant. Thus an intercomparison of the weights will yield the result:

---

[1] Figures in brackets indicate literature references at the end of this paper.

$$(M - \varrho V) - (M_x - \varrho V_x) = m \qquad (1)$$

where m is a small mass difference, read on the scale of the balance. (The acceleration of gravity, being the same in all measurements at the same location, has been cancelled from both sides of the above equality.) It is assumed in the above equation that ϱ is constant during the comparison.

The mass of the unknown is calculated from the relation

$$M_x = M + \varrho(V_x - V) - m \qquad (2)$$

which requires a knowledge of the density of air and of the difference in volume of the artifacts. By intercomparing the standard kilogram with kilogram artifacts of different volume and by conducting experiments at different values of ϱ, eq (2) may be tested. By comparing the standard with a mass equal in volume but different in surface area, one may in principle test whether additional, surface-dependent terms must be included in eq (1).

2. Experimental

2.1. Balances

Measurements at both Sandia and NBS were carried out on commercially available kilogram balances. The balances were single-pan of conventional design and each had a precision of 25–50 μg.[1] Two modifications to the balance case were made. The original glass door on the left side of the balance was replaced with one having a port which could accommodate a Dunmore-type humidity element. In addition, an annex to the balance was constructed and placed in contact with the glass door on the right side of the balance. The annex was made of metal but had glass doors and a glass floor. The floor-height of the annex was made equal to that of the balance. The annex was made large enough to accommodate the four one-kilogram weights used in any given intercomparison. Since the balance case had no room for weights in addition to whatever was on the pan, it was hoped that the annex would help minimize changes in ambient conditions as the weights were manipulated in the course of a measurement.

Manipulations were performed by an experimenter seated in front of the balance. The experimenter wore an apron of metallized Mylar to reduce the effect of his presence on the temperature of the balance. In addition, the experimenter's right hand (used for weight manipulations) was covered by an inner cotton glove and an outer surgical glove. The purpose of the gloves was to help insulate the balance and annex from changes in temperature and hu-

midity as weights were shuttled from one to the other. The same experimenter performed all the measurements reported below.

2.2. Weights

Ten different one-kilogram weights were used in the experiment. Their designations and major features are shown in table 1. The most conventional weights, B1 and D2, were used as standards. They have the desirable properties of

TABLE 1

Artifact Designation	Nominal Mass (Kg)	Volume (cm³) at 20 °C	Nominal Surface Area (cm²)
B1	1	127.385	145
D2	1	127.625	145
H1	1	337.381	270
H2	1	337.666	270
R1	1	126.395	270
R2	1	126.392	270
S1	1	126.549	660
S2	1	126.545	660
A	1	359.488	280
T	1	60.027	85

Characteristics of the 10 artifacts used in these measurements. All weights except A and T have polished stainless steel surfaces. Artifact A is made of solid aluminum alloy and Artifact T is made of solid tantalum.

single-piece stainless steel construction, knobs for ease of handling, and nearly minimum surface area. Weights H1 and H2, also of stainless steel, were designed to have a density near that of aluminum. They are hollow, right circular cylinders of minimum surface area (diameter equal to height), each having an internal center-post to lend rigidity to the end-pieces. The hollow weights are filled with helium at roughly one atmosphere pressure. The two weights R1 and R2 were constructed as companions to the hollow weights. They are solid thick-walled stainless steel tubes whose surface areas are nominally equal to those of H1 and H2. Two additional weights, S1 and S2, of solid stainless steel but with surface areas roughly twice those of the R weights were also included. The S weights are each in the form of two nested stainless steel tubes reposing on a circular, stainless steel base. A centerpost welded to the base allows easy manipulation of the S weights. The final two artifacts in the assemblage were single-piece weights of aluminum and tantalum, designated A and T. The aluminum weight, constructed of bar stock, is in the form of a right circular cylinder of minimum surface area. The tantalum weight is of single-piece construction of nearly minimum surface area with a knob for ease of handling. The aluminum and tantalum weights are the same ones as were used in the experiments reported in [1]. The weights were lifted with hand-held instruments designed for the purpose.

---

[1] The precision of the balance is determined from the experimental scatter (1 S.D.) in a set of repeated measurements of a single weight the density of which is close to that of the built-in balance weights and counterpoise. Thus the measurement of the precision of the balance is unaffected by the usual fluctuations in the density of air in the balance case.

All the weights except A and T were steam cleaned prior to the experiments reported here. The weights having stainless steel surfaces were also vapor degreased with 1,1,1-trichloroethane. No further cleaning was attempted throughout the course of the measurements reported here. All weights were dusted with a soft, lint-free brush prior to each use, however.

2.3. Measurement of Inputs to Buoyancy Equation

2.3.1. Temperature

The temperature of air inside the balance case was measured in two different ways. First, 13 thermocouples in series were disposed about the weighing chamber. Type E [4] thermocouples were chosen in order to achieve maximum sensitivity at room temperature. The reference junctions of the thermocouples were thermally anchored around a massive copper block which was itself surrounded by 8 cm of thermal insulation. The temperature of the copper block, never more than two degrees cooler than the balance temperature, was found by measuring the resistance of a capsule-type platinum resistance thermometer embedded in the center of the block. All leads were in good thermal contact with the block. The voltage developed across the thermocouples was read to $\pm 3\mu V$ with a portable potentiometer while the platinum thermometer was monitored using traditional bridge techniques [5]. The thermocouples were calibrated in the temperature region of use against platinum resistance thermometers.

A mercury-in-glass, total immersion thermometer was used as a back-up to the thermocouples. Because its range, 20 °C to 30 °C, is graduated in tenths of a degree, an observer can read the thermometer to a precision of 20 mK. However, several precautions were taken to insure the accuracy of the thermometer, which has a long history of calibration. The lag constant of the thermometer was determined in still air by measurement of its response to a temperature step of +5 °C. The resulting lag time, 185 ± 5 s, is in accord with similar measurements in the literature [6]. In an effort to reduce the lag, foil vanes were then attached to the thermometer bulb. Held in place with a thin layer of copper-impregnated vacuum grease, the vanes decreased the lag to 160 ± 5 s. A check of the modified thermometer against a platinum resistance thermometer verfied that the addition of the vanes did not change the calibration.

In addition to the lag constant, it was also desirable to measure the pressure coefficient of the thermometer arising from the elasticity of the thin-walled bulb. This was accomplished by the use of a vacuum chamber which has transparent walls and is partially filled with forepump oil. The thermometer was hung in the oil and observed with a telescope. The large heat capacity of the bath served to anchor the temperature of the thermometer during measurements. Thus the pressure coefficient of the thermometer could be determined from readings of temperature as a function of the pressure at the free surface of the oil bath. Pressure was cycled several times without indication of hysteresis. The observed coefficient, -6.9×10^{-7} °C Pa^{-1}, translates to a change in calibration of 0.012 °C between NBS and Sandia.

Of all the measurements made in the course of this experiment, temperature proved to be the most elusive. A fundamental problem, discussed in section 4.1.1, is that the balance is not in thermal equilibrium. In fact, its temperature as a function of time may be rather complex. A second problem peculiar to our particular measurements was that the thermocouple sensitivity was found to change significantly over the course of the experiment. It was discovered that the cotton insulation of the thermocouples was not robust enough to survive repeated dismantling and reassembly. As a result, a few of the thermocouples might short out unnoticed. For this reason, the thermocouple data could not be used except semiquantitatively. We therefore relied on the mercury-in-glass thermometer for the calculation of air density. In the first experiments at NBS we attached a vane to the thermometer bulb in a manner different from that described above. The resulting lag time, though not as short as the 160 s achieved later at Sandia and back at NBS, was no worse than that of the unmodified thermometer (185 s) and therefore was a minor source of error, as discussed below.

2.3.2. Pressure

Barometric pressure at both NBS and Sandia was read using aneroid barometers. The aneroids have limited pressure range so that two were required, one with a range covering air pressures near sea level and the second spanning atmospheric pressures at an altitude of $\sim$ 1600 m above sea level. The aneroid used at NBS was calibrated by the pressure calibration group of the NBS Thermophysics Division. This aneroid was also checked at ambient pressure twice daily against a cistern-type mercury manometer [7].

The aneroid used at Sandia was calibrated first by a private laboratory. In designing the experiment, it was considered prudent to recalibrate the high-altitude aneroid *in situ* at Sandia. It was felt (rightly as measurements proved) that despite hand carrying, the aneroid might change its calibration during transport to Sandia. Therefore, in addition to our mercury manometer, we took with us a piston gage and a sensitive quartz pressure transducer. The quartz transducer was used to calibrate the aneroid and was itself calibrated twice daily against the piston gage. Both the piston gage and mercury manometer data require a knowledge of the local acceleration of gravity, g, to yield an accurate pressure measurement. This number was kindly pro-

29

vided us by our hosts at Sandia. It is based on a 1976 survey of their laboratory. The calibrated quartz transducer agreed with the mercury manometer to within 30 ppm.

2.3.3. Relative Humidity

Relative humidity was read by a Dunmore-type humidity sensing element mounted in one door of the balance case. The elements used were calibrated at NBS by the Thermal Processes Division. Periodic checks of the calibration were effected by immersing the element in air above a standard salt solution [8]. The solution chosen produces an atmosphere of 43 percent relative humidity (R.H.), which was close to ambient conditions both at NBS and Sandia. The humidity sensing element has a temperature coefficient which must be taken into account but no pressure coefficient could be detected over the range of use.

2.3.4. Carbon Dioxide

The carbon dioxide (CO_2) content of the air in the balance case was tested twice daily at both NBS and Sandia. Samples of air were drawn into evacuated glass spheres, sealed with vacuum-type stopcocks, and then sent to the NBS Gas and Particulate Science Division for analysis.

2.4. Experimental Design

Weighings always followed a so-called "four-ones" pattern in which four objects of nominally equal mass are intercompared in each of the six possible combinations. A least-squares fit to the data can be obtained which then assigns the mass to each of three of the objects providing that the mass of the fourth object is known [9]. In our measurements, B1 and D2 were used in every four-ones measurement. Since the masses of both B1 and D2 are known, their sum is used in the least-squares solution. The difference in mass of B1 and D2 is also known but is not used to constrain the least-squares solution. Therefore, a comparison of the computed difference in mass with that of the accepted difference can serve as the basis of a t-test of the four-ones solutions.

Each of the six intercomparisons of the four-ones series was performed by placing weights on the single-pan balance in the following order and observing the balance indication.

1. W_1
2. W_2
3. $W_2 + \Delta$
4. $W_1 + \Delta$
5. W_1

where W_1 and W_2 are the weights to be compared and Δ is a 20 mg sensitivity weight, used throughout the measurements. Common metrological practice, often referred to as "double substitution," would require operations 1 through 4 only [10]. The fifth operation is used solely to provide a better estimate of balance zero drift. It is assumed that any changes are linear with time. Prior to measurements, small, calibrated weights were added as necessary to the four weights used in a design in order that the balance indication of all four weights be within 5 mg of equality. Added weights needed at Sandia were generally different from those at NBS because of differences in the buoyant forces acting on the kilogram artifacts at the two locations. Values of the sensitivity weight and the added weights were derived from routine calibrations traceable to the SI (International System) definition of the unit of mass. The estimated uncertainties associated with these weights are negligible with respect to the standard deviation of the balances used.

Temperature, barometric pressure and relative humidity measurements were recorded between readings 2 and 3 in the order shown above. These values were assumed to hold during all five measurements.

In addition to B1 and D2, the other two weights in each four-ones design were either R1 and H1, R2 and H2, S1 and S2, or A and T. The same pairings were maintained throughout the measurements. Each pair, other than the standard, was thus used four times—once each day. The order in which weighings proceeded for a given day was permuted from day to day so that each pair was compared once at each of the four daily periods in which mass comparisions were made. This procedure has the effect of averaging out any errors which depend on the order or the time of day in which mass comparisons are made.

One four-day sequence was carried out at NBS. The weights and ancillary equipment were then transported to Sandia where another four-day sequence was carried out. Finally, the assemblage was returned to NBS for a third four-day sequence.

3. Results

The results of the measurements are shown in figures 1–10. The three series of measurements are labelled "NBS I," "Sandia," and "NBS II" for ease of discussion. Data labelled "NBS III" and "NBS IV" refer to additional experiments, described below. The results of 4 of the four-ones series have been excluded because of obvious statistical differences with the 44 remaining measurements.

It is also of interest to plot the difference in mass between companion artifacts having the same surface areas—H1 and R1, H2 and R2, and S1 and S2. These results are shown in figures 11–13.
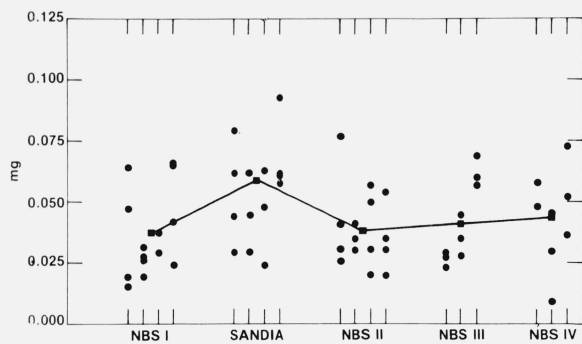
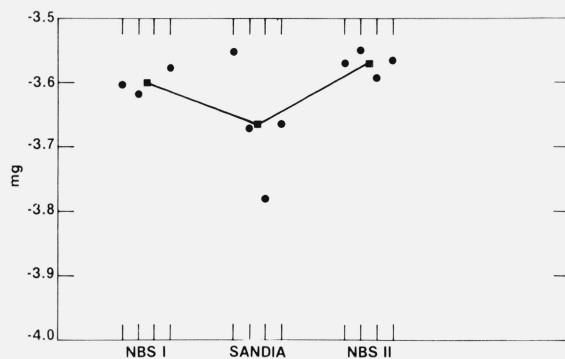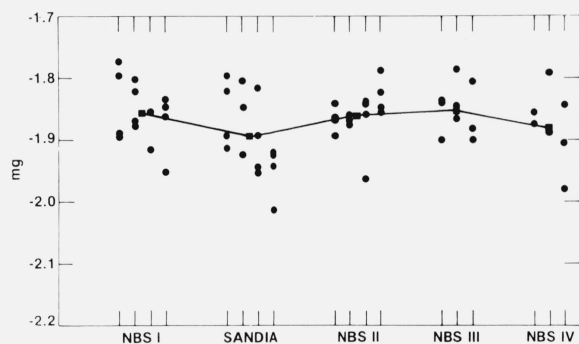FIGURE 1. *Standard deviations of least squares fits to the data.*

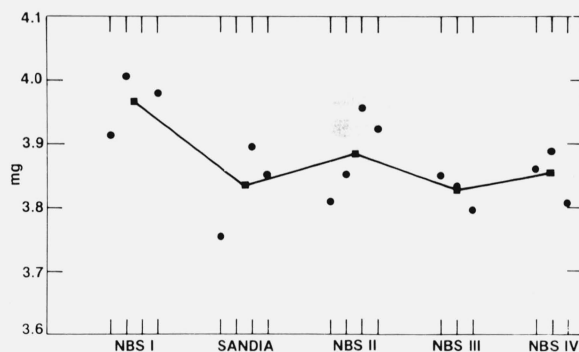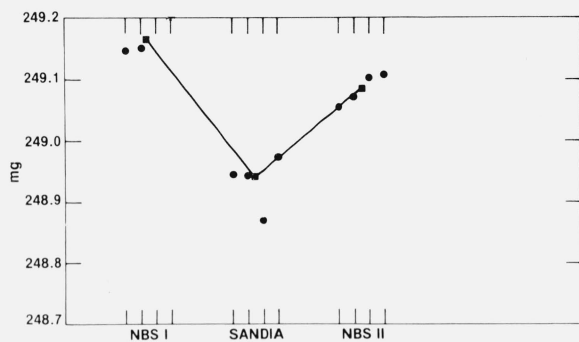FIGURE 2. *Measured differences between the mass of B1 and D2.*

FIGURE 3. *Measured differences between the mass of H1 and 1 kilogram.*

FIGURE 4. *Measured differences between the mass of R1 and 1 kilogram.*

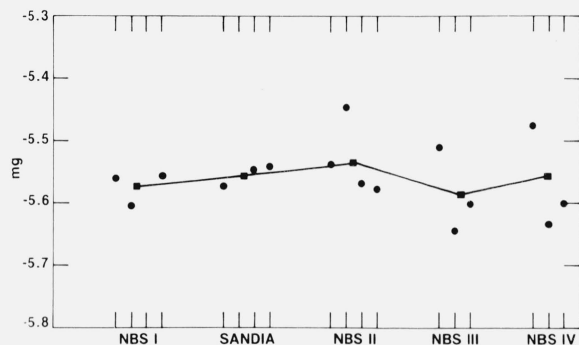FIGURE 5. *Measured differences between the mass of H2 and 1 kilogram.*

FIGURE 6. *Measured differences between the mass of R2 and 1 kilogram.*
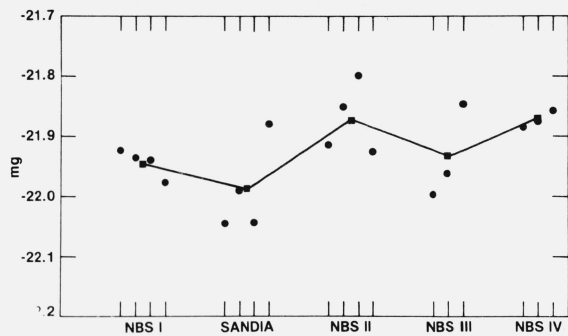
31

FIGURE 7. *Measured differences between the mass of S1 and 1 kilogram.*
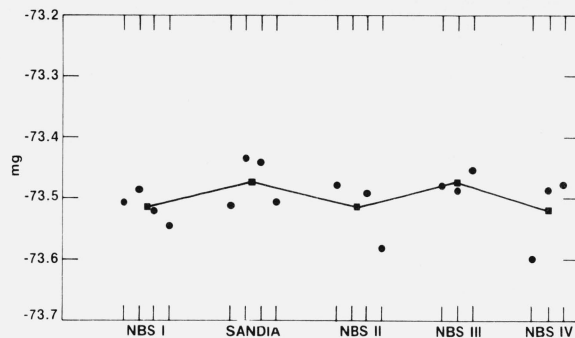

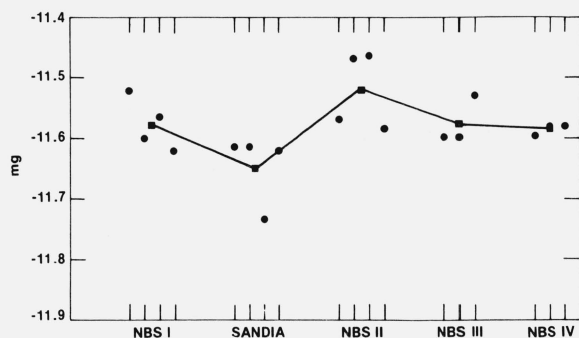FIGURE 10. *Measured differences between the mass of T and 1 kilogram.*

FIGURE 8. *Measured differences between the mass of S2 and 1 kilogram.*
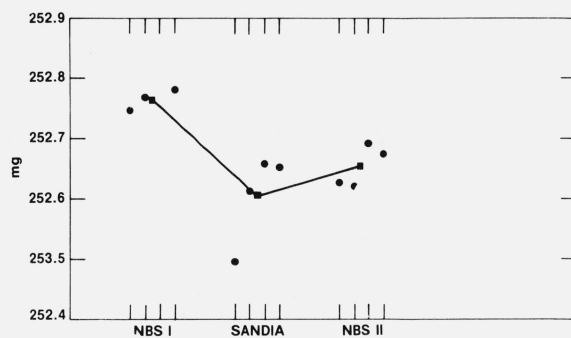

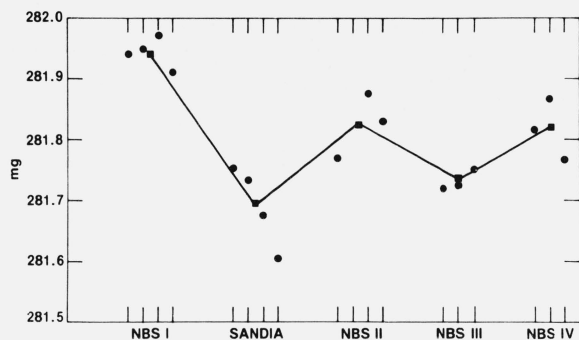FIGURE 11. *Measured differences between the mass of H1 and R1.*

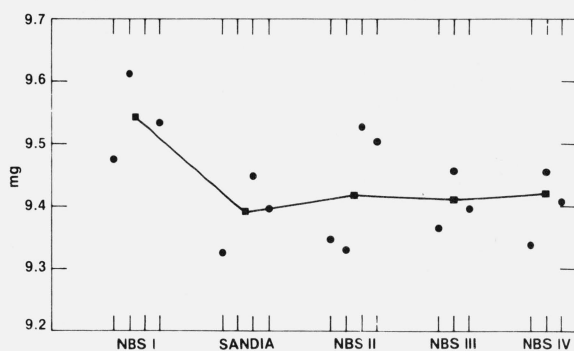FIGURE 9. *Measured differences between the mass of A and 1 kilogram.*

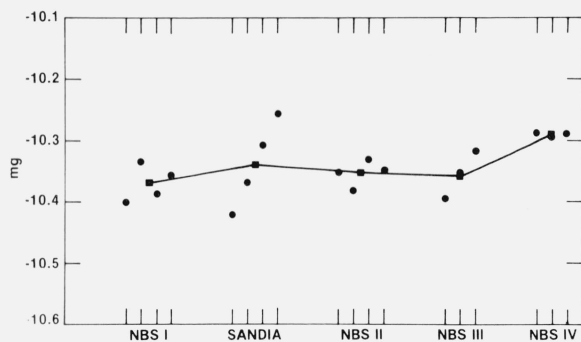FIGURE 12. *Measured differences between the mass of H2 and R2.*

32

FIGURE 13. *Measured differences between the mass of S1 and S2.*

The figures all have the same format: the abscissas, identical throughout, are ordered chronologically but spaced arbitrarily. Results obtained during a single day are plotted as vertically aligned points. Squares represent unweighted averages of each data set, lines connecting adjacent squares serving only to make trends more evident.

Table 2 shows the mean ambient conditions which obtained during the three sets of measurements. It will be noticed that the ambient temperature was lower at Sandia than at NBS. This difference reflects conditions in the two mass laboratories and was not easily altered. To test the effects of different ambient temperatures at the same location, two additional experiments were performed at NBS in a room the teperature of which could be controlled to ±0.1 °C. The experiment was first performed at ∿21 °C and then repeated at ∿23°C. The procedure was identical to that reported above with the following three exceptions: 1) the pair H1, R1 was excluded. The resulting three pairs were each measured once during each of three days; 2) the concentration of carbon dioxide in the air was not measured. A value of 0.043 percent by volume was assumed, based on a previous survey of the NBS mass laboratory; 3) the weights in the balance annex were placed on copper pads. Thermocouples, referenced to the balance chamber, were attached to the pads.

TABLE 2

	Temp (°C)	Barometric Pressure (kPa)	Relative Humidity (%)	CO₂ Concentration (% by volume)	Density of Air (mg/cm³)
NBS I	23.5	100.0	43	.053	1.169
SANDIA	21.1	83.6	40	.060	0.985
NBS II	23.2	100.3	41	.043	1.173

Columns 1–4 list the average values of the input parameters to the air-density equation. Column 5 is the average of the air densities computed at each location. Row 1 refers to the first set of weighings performed at NBS. Row 2 refers to the weighings performed at Sandia. Row 3 refers to the repetition of the weighings at NBS.

The results of these measurements, labelled "NBS III" (21 °C) and "NBS IV" (23 °C), are also shown in figures 1–13 and table 3. It was found that the temperature of the copper pads placed in the annex lagged that of the air in the balance case by ∿0.2 °C during all measurements.

TABLE 3

	Temp (°C)	Barometric Pressure (kPa)	Relative Humidity (%)	CO₂ Concentration (% by volume)	Density of Air (mg/cm³)
NBS III	20.9	101.3	30	.043*	1.199
NBS IV	23.2	99.3	28	.043*	1.164

*Assumed value

The column headings are identical to those of table 2. Row 1 refers to weighings performed at NBS in a room whose temperature was controlled at 21 °C. Row 2 refers to weighings performed in the same room with the temperature now controlled near 23 °C.

4. Discussion

4.1. Meausrement Errors

4.1.1. Temperature

The mercury-in-glass thermometer was calibrated by the NBS Temperature Measurements and Standards Division. The estimated uncertainty of their calibration is 30 mK. We checked single points on the thermometer against a calibrated platinum resistance thermometer (PRT) daily during the experiments. No significant deviations from the assigned calibrations were found. As described above, a pressure correction to the thermometer calibration was measured for the change in altitude between NBS and Sandia. Although this correction (12 mK) is less than the uncertainty in the thermometer calibration, it was nonetheless applied in the ensuing computations.

The lag time of the thermometer was measured to be approximately three minutes. The temperature rise, as measured in the balance during a four-ones series, was always about 200 mK over a period of about 45 minutes. During this time, the thermometer was read six times—once during each of the double substitutions. The effect of the thermometer lag in an environment of steadily rising temperature is that the temperature observed is the actual temperature which occurred three minutes before. The difference between observed and instantaneous temperature, about 15 mK, is nearly the same for all the measurements. The difference in time lag between the thermometer as first used at NBS and as used in the rest of the measurements is seen to

33

have a negligible systematic effect on the temperature. Lag corrections were not applied to the temperature data because of the small size of such a correction and its large relative uncertainty under actual experimental conditions.

On a time scale which is short compared to the response time of the thermometer, the temperature in the balance case is complex. This was learned by observing the thermocouples. Upon introducing one of the large-volume weights into the balance, the temperature as measured by the 13 thermocouples fell in a matter of seconds by as much as 200 mK. After about one minute the temperature of the thermocouples was seen to have risen to a slightly warmer temperature than observed before the weight was introduced. That this behavior is not observed with the artifacts of smaller volume may be a consequence of the distance of their surfaces from the thermocouples. Because the test weights are not isothermal with the balance enclosure during the measurements, the choice of the proper temperature to apply in computing the density of air is ambiguous. Although we assume an uncertainty in temperature of 30 mK, this number must itself be viewed as uncertain.

4.1.2. Pressure

The aneroid barometers used were calibrated twice daily against instruments which can read pressure with a relative uncertainty of 5×10^{-5} (mercury manometer and quartz transducer; see above). The aneroids themselves are known to deviate by as much a 0.1 mm of Hg (1 mm of Hg equals 133.3224Pa) from their calibrated value over the course of a day. Therefore the uncertainty in the pressure inside the balance case was less than 0.1 mm of Hg, a number which, we feel, approximates one standard deviation.

4.1.3. Relative Humidity

A well-behaved and well-cared-for Dunmore-type hygrometer will retain its calibration to better than 0.5 percent relative humidity for long periods of time [11]. Our elements were checked against salt solutions [8, 12] the vapor pressure of which was in the middle of the range of the humidity element. These checks established the stability of the Dunmore-type elements to 1 percent relative humidity. Temporal changes observed in the humidity sensor readings over standard salt solutions were used to estimate uncertainties in the readings. These deviations were not viewed as changes in the calibration of the elements.

4.1.4. CO_2

The carbon dioxide content of the air in the balance case was measured twice daily during the 48 four-ones series run at NBS and Sandia. These results are summarized in table 4. Since the measurement of CO_2 concentration in any given sample can be made with an uncertainty of 25 ppm in the concentration, the observed standard deviations at NBS I and Sandia represent real fluctuations of the CO_2 concentration in the balance case. As these fluctions may be observed between daily readings, we feel that the standard deviation of the fluctuations is a reasonable estimate of our uncertainty in the CO_2 concentration in the balance case. The average obtained for NBS III agrees excellently with previous measurements of ambient air at NBS. We cannot explain the significant difference between the NBS I average and the other NBS data.

TABLE 4. *Concentration of CO_2 (ppm by volume)*

	Ave	Standard Deviation of a Single Measurement
NBS I	530	60
SANDIA	600	115
NBS II	430	25

The CO_2 concentration was not measured during the final weighings at NBS which were performed in a temperature-controlled room. Instead, a value of 430 ppm by volume was assumed. We estimate an uncertainty of 100 ppm in the concentration as a result of this assumption. This uncertainty propagates as an uncertainty in ϱ of 40 ppm [2] and an uncertainty of less than 15 μg in the assignment of mass to an aluminum kilogram as calibrated against a stainless steel standard.

4.1.5. Volume of the Artifacts

The volume of each artifact was determined by hydrostatic weighing with independent mechanical checks for those with simple geometry. It is believed that all volumes are known to about 50 ppm. Of course, the volumes vary with temperature but this effect is small (<70 ppm/ °C in all cases) and, therefore, easily estimated to sufficient accuracy. Note that even a 1 percent error in ϱ could be tolerated in a hydrostatic determintion of volume to 50 ppm.

4.1.6. Air Density Equation

The use of an equation to determine the density of air entails errors apart from the instrumental inaccuracies outlined above. In a meticulous examination of the equation used in this study [2], Jones cites relative uncertainties of 50 ppm random and 50 ppm systematic independent of inaccuracies in the measurements of input parameters. These numbers treat the uncertainty in R, the ideal gas constant,

as a random error. In using the air density equation, however, the uncertainty in R becomes a systematic error in ϱ. Thus the relative uncertainties in the calculation of ϱ become 40 ppm random and 80 ppm systematic. These represent one standard deviation.

4.2. Uncertainty in the Buoyancy Correction

We may now calculate the uncertainty expected in applying buoyancy corrections to our weighings. There are two uncertainties of interest: random, which introduces scatter in the measurements, and systematic, which introduces errors in the average values obtained. In particular, we are concerned with an estimate of the maximum expected difference between mass measurements of the same artifact at NBS and at Sandia due to known systematic uncertainties in the buoyancy correction.

Table 5 summarizes the random uncertainties expected in the calculation of the density of air, ϱ. These will lead to random uncertainties in the buoyancy corrections of 50 μg for A, H1, and H2; 15 μg for T; and virtually zero for R1, R2, S1 and S2. These numbers assume usual laboratory conditions at NBS. At Sandia, the random uncertainties are calculated to be 20 percent smaller bacause the buoyancy correction is itself 20 percent smaller than at NBS.

TABLE 5

Source	Estimated random uncertainty in ϱ (ppm, 1 std. dev.)
Air-density equation	30
Temperature*	100
Barometric pressure	130
Relative humidity	90
CO_2 concentration	25
Combined (by quadrature)	190

*See text for discussion of temperature measurement uncertainties.

The relatively smaller magnitude of the buoyancy correction at Sandia compared with that at NBS leads to the possibility of systematic differences between masses measured at the two locations. These descrepancies arise from systematic errors in the calculation of ϱ as well as from errors in ΔV, the assignment of volume difference between an unknown weight and the standard. Table 6 details the known sources of systematic uncertainty in ϱ, while table 7 indicates the resulting uncertainties in comparing the mass of an artifact as measured at NBS with the mass of the same artifact as measured at Sandia.

Although not indicated in table 7, the sign of systematic errors due to ϱ is opposite for weights which are more dense and less dense than the standards.

TABLE 6

Source	Estimated systematic uncertainty in ϱ (ppm, 1 std. dev.)
Air-density equation	80
Barometric pressure	30
Temperature*	40
Relative humidity	50
CO_2 concentration	10
Combined (by addition)	210

*See text for discussion of temperature measurement uncertainties.

TABLE 7. *Estimate, based on table 6, of systematic differences, in micrograms, in the mass of artifacts as measured at NBS and at Sandia against stainless steel standards.*

Source of uncertainty	Artifact		
	A1, H1, H2	R1, R2, S1, S2	T
ϱ	10	0	3
ΔV	4	2	1
Combined	14μg	2μg	4μg

Thus the maximum systematic difference to be expected, at a level of one standard deviation, i. e. in measurements of H1, H2 and A, is smaller than the standard deviation of the balances used in the measurement. It should be emphasized that the only systematic effects considered in table 7 are those associated with applying buoyancy corrections to the data obtained from readings of the balance.

In order to evaluate the results, we must establish a criterion by which to assess the significance of any discrepancies observed among sets of data. Referring to figure 1, one sees that the average standard deviations of all four NBS data sets are roughly the same and equal to about 40 μg. The average standard deviation at Sandia was 60 μg. Using these numbers to define the experimental standard deviations at NBS and Sandia, we can calculate by well known techniques [14] whether the means of different sets of data differ at the 0.05 level of significance. These calculations depend on the number of independent measurements in the data sets being compared. Table 8 is a compendium of the various statistical conditions pertaining when one compares results which are plotted in figures 3–10.

At the level of significance chosen, there definitely remain systematic differences in the masses of the same object computed at different locations or times. Note that some of these systematic differences occur in objects whose volume is nominally the same as that of the standards, a situation which is nearly insensitive to systematic errors in buoyancy correction.

35

TABLE 8

Number of independent data sets used to calculate the mean			Maximum difference at 0.05 level of significance
at NBS	at NBS	at Sandia	
4	4	—	34 μg
4	3	—	37
3	3	—	39
4	—	4	43
3	—	4	45
4	—	3	48
3	—	3	50

To compare averages in figures 3–10, find the number of measurements in each data set from which the average was computed. Matching these numbers to the corresponding horizontal line of the table allows one to find the maximum difference expected at the 0.05 level of significance (95% confidence level).

Certain trends may be inferred from the systematic differences among the data sets. In general, the extremal values of the various mass determinations were found at NBS I and Sandia. Computed mass values obtained at NBS III and NBS IV agree well with Sandia values while NBS II values fall between those of NBS I and the others. These features are unchanged if the data are reanalyzed using direct comparison with the standards instead of a least squares solution to a four-ones series.

Specifically, let us compare the Sandia results with those of the other series. These comparisons are displayed in table 9. It is striking that 70 percent of the numbers displayed are negative. The tantalum weight as well as R2 and possibly H2 are the only weights immune to the negative systematic difference. Many of the differences, when looked at alone, are within reasonable expectations as calculated in the preceding paragraphs. When taken as an ensemble, however, the systematic behavior is apparent.

TABLE 9

Weight	S-NBSI	S-NBSII	S-NBSIII	S-NBSIV
H1	−.227 mg	−.145 mg	—	—
R1	−.066	−.096	—	—
H2	−.113	−.033	+.026 mg	.000 mg
R2	+.017	−.023	+.029	+.014
S1	−.043	−.116	−.055	−.116
S2	−.072	−.132	−.075	−.068
A	−.251	−.134	−.040	−.125
T	+.041	+.043	−.001	+.047

Differences (in milligrams) between masses of weights determined at Sandia (S) and at NBS.

One may now ask whether this systematic behavior can be explained by buoyancy effects. Table 10 makes clear that buoyancy cannot be the explanation. Were unexpected buoyancy effects to be seriously considered, one should see a clear dependence of the systematic effects on volume. but, while the differences between NBS I and Sandia data are consistent with a buoyancy hypothesis, the other data are not.

TABLE 10

$V_x - V_{STD}$	S-NBSI	S-NBSII	S-NBSIII	S-NBSIV
230 cm^3	−.251 mg	−.134 mg	−.040 mg	−.125 mg
210	−.170	−.089	+.026	.000
−1	−.041	−.092	−.034	−.063
−68	+.041	+.043	−.001	+.047

Differences between weights measured at Sandia (S) and at NBS. These differences, measured in milligrams, are tabulated as a function of the difference between the volume of the weights in question and the standards B1 and D2.

The behavior of S1 and S2 (figs. 7 and 8) suggests that surface effects, related to temperature, may play a role in the measurements. In fact, when the difference in mass of S1 and S2 is plotted as in figure 13, there remain no significant discrepancies as a function of place or time. Similar graphs of H1,R1 and H2,R2 (figs. 11 and 12) indicate a significant difference between the measurements of NBS I and the remainder of the data. We have no way to eliminate surface-related effects in the measurements of A and T.

The appearance of surface effects is not likely to be due to moisture. A simple calculation indicates that about six monolayers of water would have to be removed from the standards and S1 and S2 to account for the systematic differences observed between measurements at NBS and Sandia. The data of Kochsiek on the moisture content of stainless steel surfaces [13] render this possibility untenable.

It seems to us likely that the cause of most of the systematic scatter in the data is the absence of thermal equilibrium between the artifact weights and the balance. This absence of equilibrium may manifest itself as a force which depends qualitatively on the shape or surface area of the weights. Such effects have been observed in small weights [15]. In addition, the buoyancy correction assumes equilibrium conditions. The measurements designated NBS III were an attempt to duplicate the thermal environment of Sandia as nearly as possible. To this end measurements were conducted at 21 °C in a temperature controlled room. Table 4 suggests that the duplication of the Sandia thermal conditions did come closer than the other NBS measurements to duplicating the Sandia data. Nevertheless, none of the measurements were done at thermal equilibrium conditions if one considers 0.2 °C fluctuations as significant.

Clearly it is desirable to perform the above measurements under isothermal conditions. A thermostated balance enclosure and a weight-changer which will accommodate low-density kilograms were already under development before the measurements reported above were undertaken. These modifications, when completed, will permit controlled study of the effects reported above.

It should be emphasized that the largest systematic differences observed are more than a factor of five smaller than those which occasioned the publication of [1].

An additional, unexpected result deserves mention. Our data show that unpolished aluminum bar stock is well-behaved as a weight. That is, the least-squares mass solutions to four-ones weighing series which contained A as one of the weights showed consistently lower standard deviations than four-ones solutions of all-stainless steel weights.

It has been suggested that a pair of weights having nominally equal masses and surface areas but very different volumes be used to determine air density in a balance [3]. The combinations H1,R1 and H2,R2 are such pairs. In particular, we have measured the difference in mass of H2 and R2 thirteen times at NBS and three times at Sandia. The standard deviation of the NBS differences is 87 μg. These measurements were taken over a period of five months. The observed standard deviation may be taken as an indication of how well the two-artifact method of inferring the buoyancy of air agrees with the method actually used in our measurements. The disagreement indicates a random uncertainty (1 standard deviation) in the application of the two-artifact method of $\sim$300 ppm—this is consistent with the results of other experiments [3,16]. This uncertainty is also, to an unknown extent, subject to the systematic effects which we have discussed at length above. In spite of this, however, two observations may be made: 1) The data taken at Sandia do not differ significantly from the NBS results and 2) the 300 ppm uncertainty in the measurement of ϱ by the two-artifact method is consistent with the minimum random uncertainty expected for use of the air density equation with state-of-the-art measurement of input parameters [2].

5. Conclusion

1. Five groups of measurements of the mass of an aluminum and a tantalum kilogram against stainless steel standards were carried out over a period of several months. Four groups of measurements were made at NBS, Gaithersburg and one at Sandia Laboratories in Albuquerque. While the groups of data exhibit significant differences amongst them of the type reported in [1], the magnitude of these discrepancies is a factor of five less than had previously been observed by the author of [1]. We remain unable to reproduce or satisfactorily explain the earlier results.

2. The results using weights with purposely enhanced surface area and weights of stainless steel with artificially low density indicate that surface effects likely play a role in our observed discrepancies.

3. The systematic discrepancies which are present in our results cannot be explained by buoyancy effects since these discrepancies exist between objects which have nearly identical volume. Our hypothesis is that the observed behavior is due to the weights not being in sufficiently good thermal equilibrium with the balance. This hypothesis will be explored using apparatus now under construction.

4. In the present experiment, the procedures followed to tie measurements of pressure, temperature, and relative humidity to absolute standards were the most rigorous which are likely to be used for routine mass calibrations of high precision. In addition, extraordinary precautions (short of remote control) were taken to reduce the effect of operator proximity on the measurements. We believe our results demonstrate the systematic errors which may be expected even under these circumstances.

---

6. References

[1] Pontius, P. E., Science **190**, 379 (1975)

[2] Jones, F. E., J. Res. Nat. Bur. Stand. (U.S.) **83**, 419 (1978)

[3] Koch, W. F., Davis, R. S., Bower, V. E., J. Res. Nat. Bur. Stand. (U.S.) **83**, 407 (1978)

[4] Powell, R. L., Hall, W. J., Hyink, C. H., Jr., Sparks, L. L., Burns, G. W., Scroger, M. G., Plumb, H. H., NBS Monograph 125—Thermocouple Reference Tables Based on the IPTS-68 (Nat. Bur. Stand. (U.S.), March 1974) pp. 90–102

[5] Riddle, J. L., Furukawa, G. T., Plumb, H. H., NBS Monograph 126—Platinum Resistance Thermometry (Nat. Bur. Stand. (U.S.), April 1973)

[6] Harper, D. R., III, Bul Bur. Stand. **8**, 659 (1912)

[7] Brombacher, W. G., Johnson, D. P., Cross, J. L., NBS Monograph 8—Mercury Barometers and Manometers (Nat. Bur. Stand. (U.S.), August 1964)

[8] Stokes, R. H., Robinson, R. A., Ind. and Eng. Chem., Analyt. Ed. **41**, 2013 (1949)

[9] Cameron, J. M., Croarkin, M. C., Raybold, R. C., NBS Tech. Note 952—Designs for the Calibration of Standards of Mass (Nat. Bur. Stand. (U.S.), June 1977)

[10] Pontius, P. E., Cameron, J. M., NBS Special Publication 300—Precision Measurement and Calibration, Ku, H. H., Ed., Vol. I—Statistical Concepts and Procedures (Nat. Bur. Stand. (U.S.), Feb. 1969) pp. 1-20

[11] Handegord, G. O., Hedlin, C. P., Trofimenkoff, F. N., Humidity and Moisture Measurement and Control in Science and Industry, Wesler, A., Ed.-in-chief, *Vol. I—Principles and Methods of Measuring Humidity in Gases*, Ruskin, R. E., Ed. (Reinhold Publishing Corp., N.Y. 1965) pp. 265-272

[12] Meites, L., *Handbook of Analytical Chemistry* (McGraw-Hill Book Co., Inc., N.Y. 1963) p. 3-29

[13] Kochsiek, M., PTB-Mitteilungen **87,** 478 (1975)

[14] Natrella, M. G., NBS Handbook 91—Experimental Statistics (Nat. Bur. Stand. (U.S.), Aug. 1963), Chapter 3.

[15] Blade, E., Ind. and Eng. Chem., Analyt. Ed. **12,** 330 (1943)

[16] Toropin, S. E., Snegov, V. C., Izmeritelnaya Tekhnika **12,** 75 (1975). English trans.: Meas. Tech. **19,** 1847 (1976)